

小山工業高等専門学校 電気電子創造工学科

平成 30 年度 卒業論文

広域センサネットワークを用いた

トラッキングサービスにおける

サービスに不要なプライバシーデータ流出比の定量評価

Quantitative Evaluation of Ability for Privacy Protection Using Simulator of
Local Production for Local Consumption Type Networks
in Wide Area Sensor Tracking Service

電気電子創造工学科

田村 峻

Department of Innovative Electrical Electronic Engineering
Ryo TAMURA

指導教員 干川 尚人

Naoto HOSHIKAWA

学科長	指導教員

1. 背景

近年の Internet of Things (IoT) 技術の発展とオープン化により、町中にある様々な機器がサービスへ利用可能になると予想されている。さらに高精度な画像処理技術が合わさることで、センサネットワークを利用した情報処理サービスはより広域かつ高精度になると期待される。一方でこのようなサービスでは無差別なセンシングが頻繁に行われ、サービスには不要なプライバシーデータ (Non-related Privacy Data, 以下 NPD と記載) も収集してしまう課題がある。顔などの生体関連のプライバシーデータ提供を不安に感じる利用者は特に多い[1]ため、もし有用なサービスが提案されても社会実装に理解が得られない可能性がある。この課題に対し、本研究チームは広域センサによる見守りサービスにおける NPD 流出を抑制できる地産地消型ネットワークアーキテクチャ (Local Production for Local Consumption Network Architecture, 以下 LPLC モデルと記載) を提案している[2]。本稿では広域見守りサービスにおける NPD 流出比を算出する数理モデルを示し、それを再現するシミュレータを示す。またそのシミュレータを用いて、LPLC モデルのプライバシー保護能力を定量評価する。

2. 数理モデル提案

2.1. LPLC のモデル化

LPLC の動作モデル図を図 1 に示す.

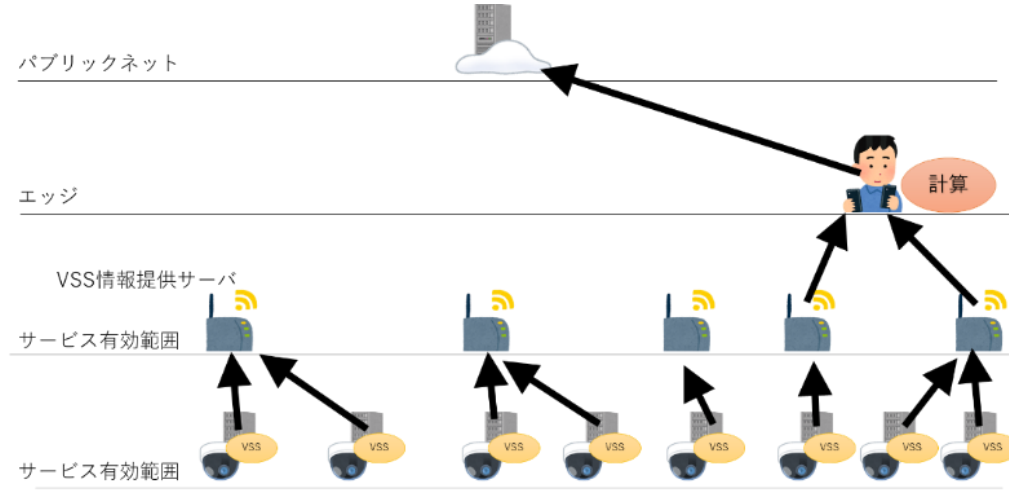


図 1.LPLC の動作モデル図

LPLC モデルでは、追跡対象者の持つ計算ノード(以下 LPLC ノードと記載)がルータアクセス可能範囲内にある時、ルータと同セグメントにあるセンサデータを処理する。そのため全てのデータを収集するクラウドモデルと比べて、NPD 流出を制御できる。なお、ここでは計算ノードにデータが流入することを流出と定義する。サービス領域を内包するように長方形(以下マップと記載)を置き、それを外包するように正方形のセルを敷き詰める。それぞれのセルが持つ対象者の有無、センサ有効範囲内といった情報を各行列の要素に対応させモデル化する。まず n 番目のカメラ集合を定義する。この集合は、各セルがカメラのセンサ有効範囲内か否かを格納した行列 $C_n(t)$ と、マップ上でのカメラの撮影方向 $up, down, right, left$ の逆 $C_{dn}(t)$ を要素に持つ。次に n 番目のルータの集合 R_n を定義する。これは属するカメラ集合の集合 RC_n とアクセス可能範囲行列 A_n を要素に持つ。ここでルータ位置とアクセス可能距離をそれぞれ $(x, y), r$ とすると、 A_n は次のように定義できる。

$$[A_n]_{ij} = \begin{cases} 1 & \text{if } x-r \leq i \leq x+r \text{ and } y-r \leq j \leq y+r \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

最後に n 人目の対象者の集合を定義する。この集合は各セルでの対象者の有無を格納した行列 $P_n(t)$ と対象者の向いている方向を表現する行列 $P_{upn}(t)$, $P_{downn}(t)$, $P_{rightn}(t)$, $P_{leftn}(t)$ を要素に持つ。例として、 $P_{upn}(t)_{ij}$ が 1 であるとは、 n 人目の対象者が時刻 t の時にセル ij でマップ上方を向いていたことを示している。それぞれの行列の定義を下に示す。また、 $P_n(t)$ の総和を $P(t)$ と定義する。方向を示す行列も同様である。非対象者集合 $N_n(t)$ も同様に定義する。

$$\begin{aligned}
[P_{up_n}(t)]_{ij} &= \begin{cases} 1 & \text{if } P_n(t)_{ij} \cdot P_n(t-1)_{(i+1)j} \neq 0 \\ 0 & \text{otherwise} \end{cases} \\
[P_{down_n}(t)]_{ij} &= \begin{cases} 1 & \text{if } P_n(t)_{ij} \cdot P_n(t-1)_{(i-1)j} \neq 0 \\ 0 & \text{otherwise} \end{cases} \\
[P_{right_n}(t)]_{ij} &= \begin{cases} 1 & \text{if } P_n(t)_{ij} \cdot P_n(t-1)_{i(j-1)} \neq 0 \\ 0 & \text{otherwise} \end{cases} \\
[P_{left_n}(t)]_{ij} &= \begin{cases} 1 & \text{if } P_n(t)_{ij} \cdot P_n(t-1)_{i(j+1)} \neq 0 \\ 0 & \text{otherwise} \end{cases}
\end{aligned} \tag{2}$$

サービス開始時間を t_0 ，終了時間を t_n とすると，NPD 流出比は次式のようになる．

$$\int_{t_0}^{t_n} \sum_{n \in R} \left\{ \delta(P(t), A_n) \frac{S_n(N(t))}{S_n(P(t)) + S_n(N(t))} \right\} dt$$

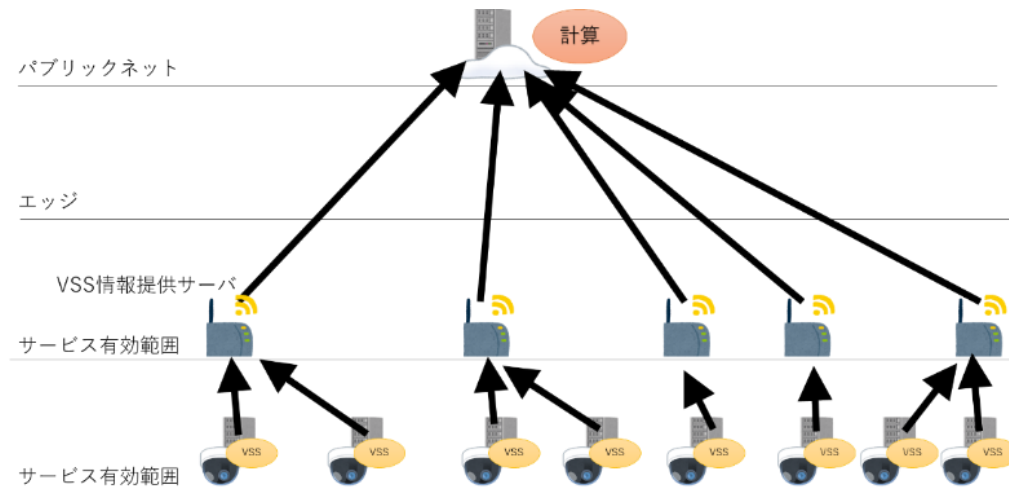
なお， $S_n(P(t))$ は n 番目のルータが時刻 t に対象者全体のセンシング量を返す関数， $\delta(P(t), A_n)$ は n 番目のルータのアクセス可能範囲内に時刻 t の対象者が存在するか否かを返す関数である．

$$S_n(P(t)) = \sum_{n \in RC_n} P_{C_{d_n}(t)}(t) \cdot C_n(t)$$

$$\delta(P(t), A_n) = \begin{cases} 1 & \text{If } Tr(P(t) \cdot {}^tA_n) \neq 0 \\ 0 & \text{otherwise} \end{cases}$$

2.2. クラウドサービスのモデル化

クラウドモデルの動作モデル図を図に示す.



図クラウドサービスの動作モデル図

既存サービスモデルは主にクラウドサービスを使って実現されている[3][4]. クラウドサービスではあらゆるデータをクラウド上にあるサーバに送信し, 画像処理はそのサーバ上で行っている. よって LPLC と違い収集するデータはセンシングされたデータ全てとなる. クラウドモデルの NPD 比は次のような式で表される.

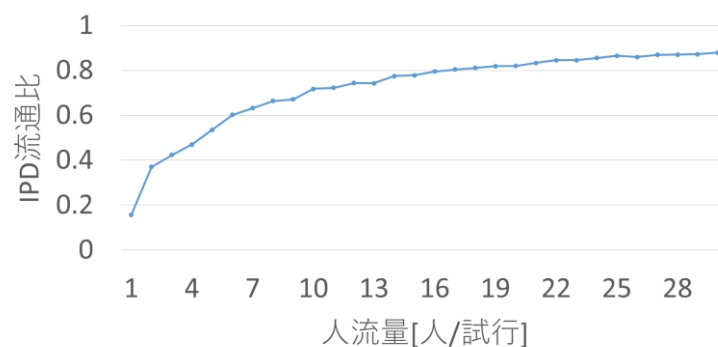
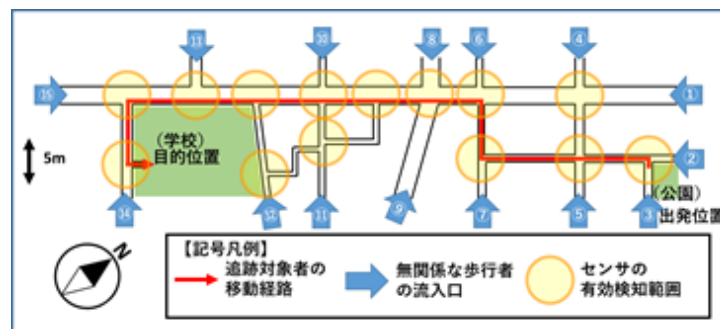
$$\int_{t_0}^{t_n} \sum_{n \in R} \frac{S_n(N(t))}{S_n(P(t)) + S_n(N(t))} dt$$

3. シミュレーション実験

3.1. 提案手法

3.1.1. プロトタイプシミュレータについて

以前簡易的なクラウドサービスのシミュレータ[5 干川先生のやつ?]を作成し、クラウドモデルの NPD 比について検討を行った。このシミュレーションでは実在する小学校の周囲をモデルにサービス領域を作成(図)し、サービス内容は「公園から学校までの見守り」がクラウドサービスで提供されるとした。見守り対象者(以下対象者と記載)を児童とし、そうでない人物を見守り非対象者(以下非対象者と記載)とした。カメラは全ての交差点に設置し、このシミュレーションではカメラの方向を考えない(数式でフォローを入れる?)。人の歩行速度を 1.25[m/sec]、カメラの動作周期を $t=1[s]$ 、セル 1 辺を 1.25[m] とした。また、追跡対象者は赤線に示す固定のルートをとる、非対象者は青矢印に示す 15 カ所の流入口から 1 つ選びそこからのルートを立ち止まることなく進む。非対象者が増減した際の NPD 流出比を図に示す。このシミュレータから、人が快適に移動できると言われる 00%未満であっても NPD 流出比は 00%を超えていることがわかった。

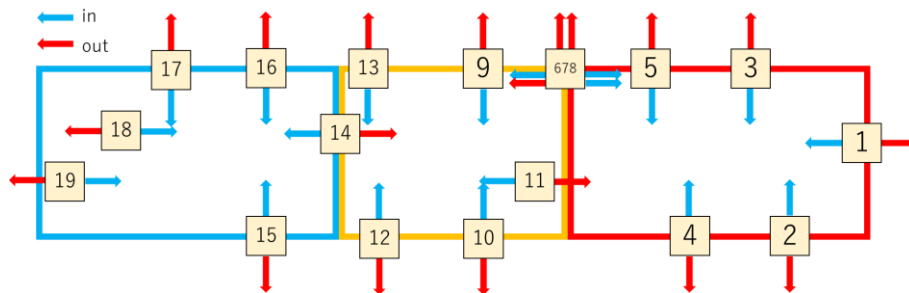


3.1.2. オープンデータについて

プロトタイプでは見守り対象者の動きや人数について自由に設定したため、現実的な NPD 流出比であるかは疑問が残った．そこで本稿では G 空間情報センターの公表している 松江駅構内人流センサデータ（西日本旅客鉄道株式会社）[6]（以下オープンデータ）を用いることで、より現実 に即した NPD 流出比を得られるようにした．オープンデータはセンサ 19 個ごとに「センサ ID, 日付, 時刻, In, Out, In 累計, Out 累計」が定義されている．また、センシング結果は 1 分毎の合計が記録されている．センサ位置, in/out データは公開されている次の画像(図)に従う．ここでは 2018 年 6 月, 平日の 5 時から 24 時のデータを扱う．なお, 人流センサ 15 番は動作を確認できなかったため扱っていない．



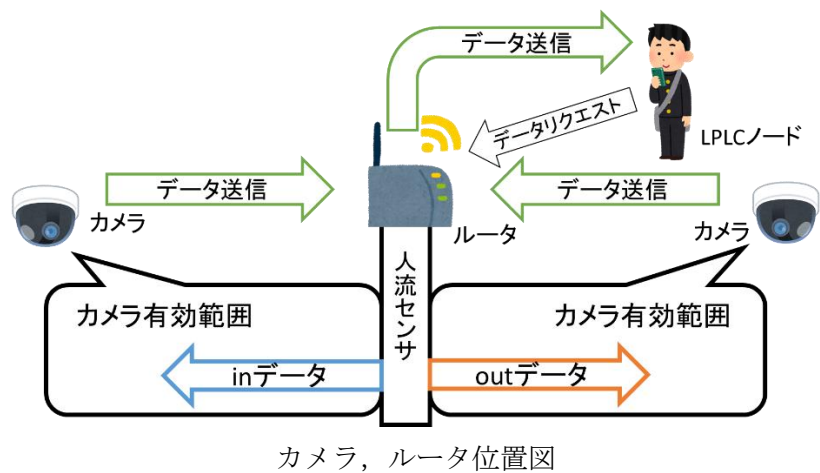
センサ設置位置図



センサ in/out 定義図

3.1.3. シミュレーション方式

オープンデータはあくまで人流センサデータであるため、このままではプライバシーデータとして扱えない．そこで in/out 人流のそれぞれを撮影する仮想のカメラを設定し、人流センサデータをカメラデータへ変換する機構を作成した(図)．カメラは RD-CA213 AHD カメラ 220 万画素バレット型赤外線搭載屋外カメラ[7]を想定し、同カメラを販売している会社の公表しているデータ[8]からセンシング可能距離を 15m とした．また LPLC モデルにはカメラの近くにルータを設置する必要があるため、オープンデータにおける人流センサと同じ位置にルータを設定した．



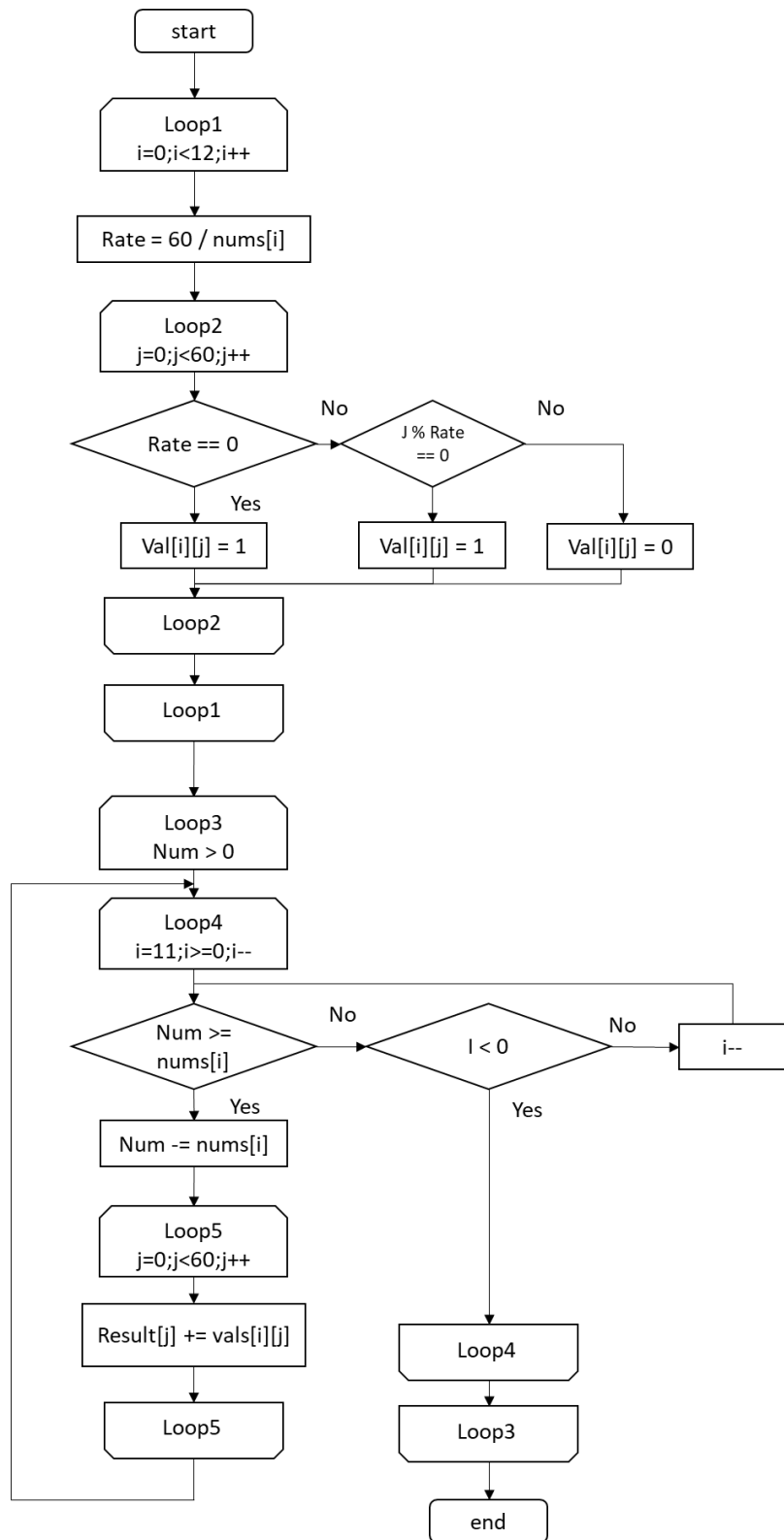
本稿ではカメラのセンシング周期を 1 秒とする．オープンデータは 1 分毎に更新されるので、次に示すアルゴリズムを用いてオープンデータを 1 秒毎のデータへ変換した．

表 1．アルゴリズムに用いる変数の定義

60 の約数	$nums[12]=\{1,2,3,4,5,6,10,12,15,20,30,60\}$
約数毎の 60 分割	$vals[12][60]$
入力データ	$input$
出力データ	$result[60]$

$vals[i]$ には $nums[i]$ を 60 分割したデータが入っている．例えば 20 の $vals$ には $\{1,0,0,1,0,0,1,\dots,1,0,0\}$ と 1,0,0 が 20 回繰り返されるデータを格納している．このように $vals$ には 60 個の要素の合計が $nums$ となりつつ、数字が偏らないように格納されている． $input$ が 0 になるまで次の処理を繰り返すことで $result$ に $input$ を 60 分割した数が格納される．

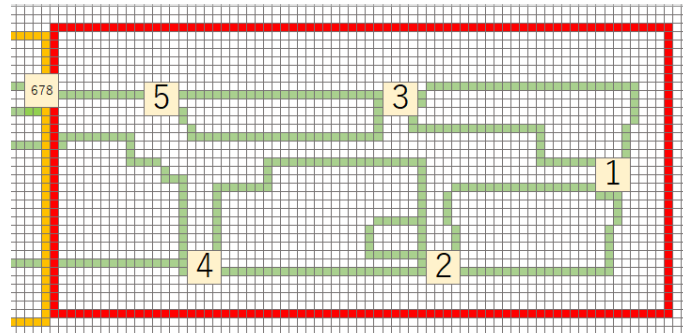
- ① $input$ より小さい中で、最も大きい $nums$ を選ぶ．選ばれた $nums$ を仮に $nums[i]$ とする．
- ② $input$ から $nums[i]$ を引き、 $result$ に $vals[i]$ を足し合わせる．



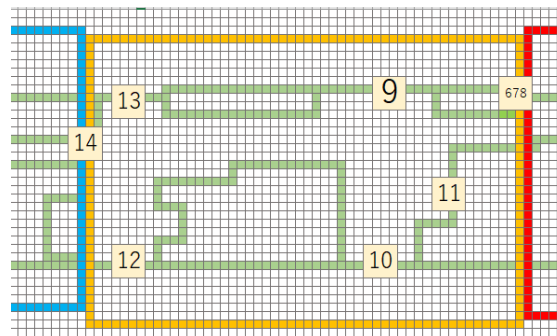
3.2. シミュレータの設計

3.2.1. サービスモデル

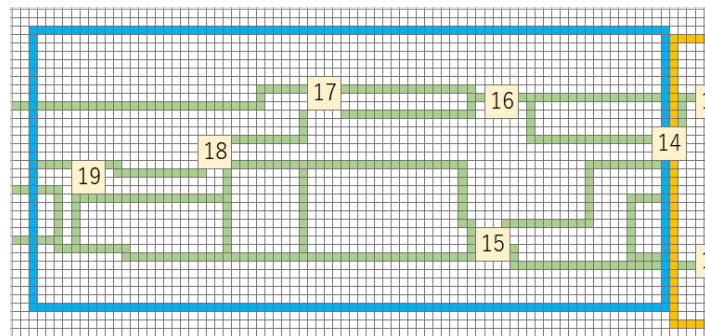
ここでは広域見守りサービスの例として、松江駅登校者見守りサービスを設定する．サービス領域はオープンデータで示された領域（図）の内，壁に囲まれていない部分とし，サービス内容を「駅を利用する高校生の登校見守り」とした．サービス時間は曜日や学校に依存しにくい登校を想定し，7時から8時とした．対象者は松江駅を利用して登校する高校生とし，そうでない人物を非対象者とした．オープンデータ全てを非対象者のデータとし，対象者のデータについては次の項で述べる．カメラの位置，方向は図で示したとおりである．人の歩行速度を $1.25[\text{m/sec}]$ ，カメラの動作周期を $t=1[\text{s}]$ ，セル1辺を $1.25[\text{m}]$ とした．モデル化された JR 松江駅を図に示す．サービス領域外を灰色で，センシング可能範囲を緑色で可視化したものを図に示す．



(a) シャミネ

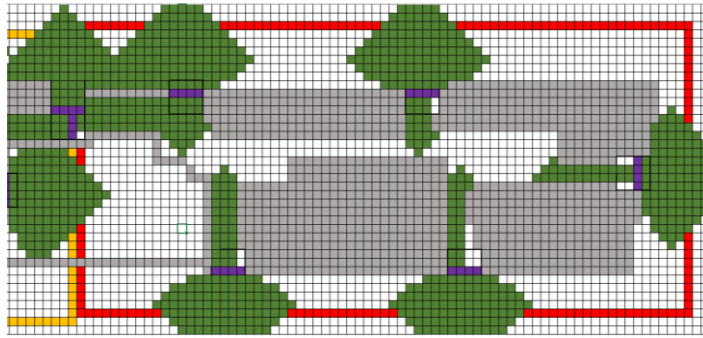


(b)

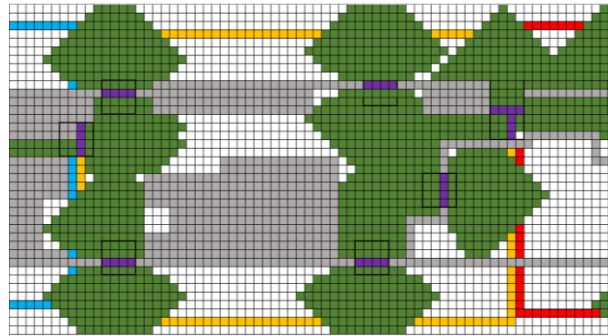


(c)

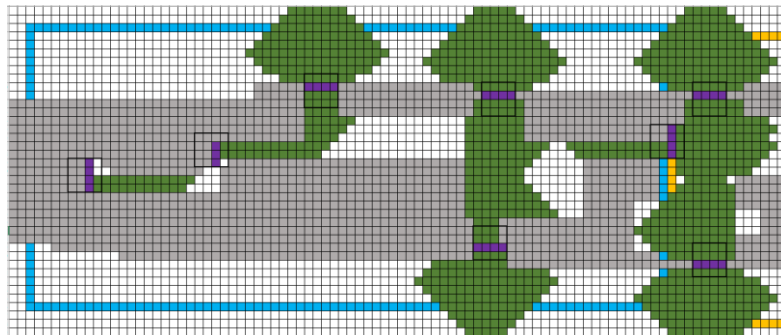
図. JR 松江駅モデル化図



(a)



(b)



(c)

図. 監視範囲図

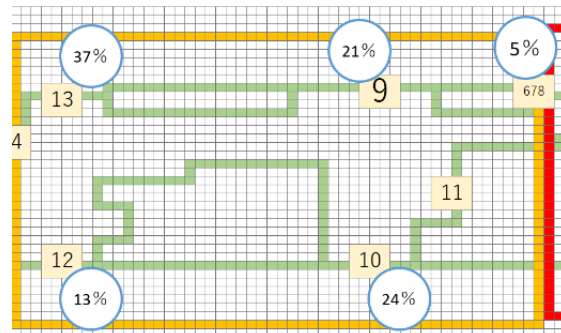
3.2.2. 対象者設定

サービス時間の人流センサデータ数の一日当たりの平均を表に示す。「センサ通過のべ人数」とは in/out データ両方を加算した値である。

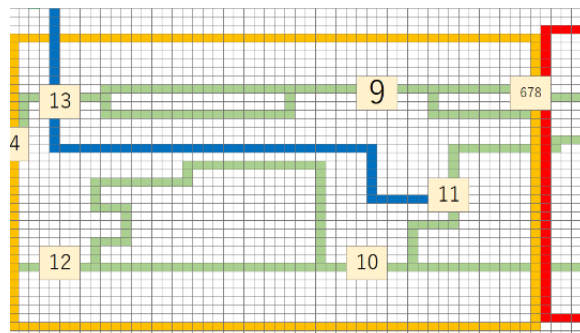
表：登校時人流センサ毎の総データ量

センサ番号	1	2	3	4	5
センサ通過のべ人数	0	0	2	0	2
センサ番号	6	7	8	9	10
センサ通過のべ人数	94	14	68	452	514
センサ番号	11	12	13	14	15
センサ通過のべ人数	1181	277	811	0	×
センサ番号	16	17	18	19	総和
センサ通過のべ人数	0	0	0	0	3415

ある人物が駅を通過する際には少なくとも駅への流入時と流出時の 2 つのデータを生成すると考えられる。この仮定と表から、一日当たりの 7 時から 8 時の駅利用者をおおよそ 1700 人分とする。鉄道利用者に占める学生の割合は全体の 7%程度[9]であるから、対象者の最大人数を全体の 10%程度とし、200 人と設定した。対象者が駅を通過する際に使う流入口は人流センサ 11 番(改札)で固定とし、流出口は物理的に駅の外側に接している人流センサから表に従い確率で選択した。利用される流出口の利用率を図に示す、人流センサ 11 番から流入し、人流センサ 13 番から流出する際の対象者のルートを図に青色で示す。



流出口使用割合



対象者移動ルート例

3.3. シミュレーション内容

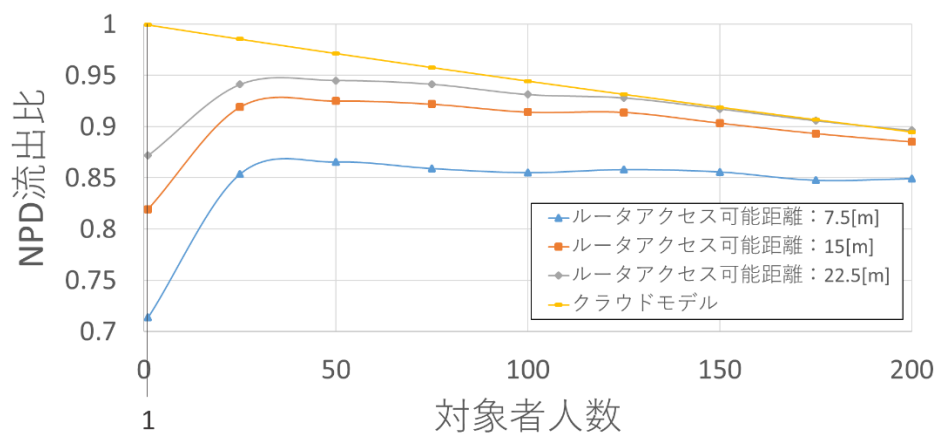
長崎

(福岡)ルータアクセス可能距離の値を設定し、対象者数を1人に加え、25人から200人まで25人刻みで変化させシミュレーションを行った。また、対象者を1人とした時にルータアクセス可能距離を変化させた場合のNPD流出比をシミュレーションした。

3.4. シミュレーション結果

長崎

ルータアクセス可能距離毎の対象者人数とNPD流出比の関係、およびクラウドモデルのNPD流出比を図にまとめた。



シミュレーション結果

表

NPD 流出比が小さいほどプライバシー保護能力が高いといえるので、ルータアクセス可能距離に依らず、クラウドモデルよりも LPLC モデルの方が優れている。

4. 考察

ルータアクセス可能距離がセンサ有効距離より短いと追跡サービスとして不十分であるが、反対にアクセス可能距離があまりに長いと NPD センシング過多の原因となる。よって、センサ有効範囲を内包しつつルータアクセス可能距離を短くすることにより、プライバシーに配慮した追跡サービスが実現できる。また図から、対象者数が少ないときの LPLC モデルのプライバシー保護能力が特に優れている。そのため対象者人数に対して LPLC ノードの数を減らすことが出来れば、より NPD 流出比を抑制できる。

センシング可能範囲と NPD 収集量の間にはトレードオフの関係がある。ただ NPD 比の削減のみを目指し続けると、そもそもセンシングをしなければよいという支離滅裂な回答が得られかねない。そこでサービス領域カバー率という指標を用意する。これはサービス領

域の面積の内何%がセンシング可能であることを示す指標である。これが 100%に近いほど、見守りサービスとして十分であると言える。本シミュレータは図と図より、サービス領域カバー率は 46.1%であった。

5. まとめ

本稿では LPLC モデル及びクラウドモデルの NPD 比を数理モデルで表現し、シミュレータから LPLC モデルはクラウドモデルと比べプライバシー保護能力に優れていると定量評価した。今後はより多人数を見守る際の NPD 比の削減方法及び LPLC モデルの実証実験等を行い、LPLC モデルについての研究を進めていく。

謝辞

- ・ G 空間情報センターの人
- ・ 渡邊先生

参考文献

- [1] 総務省. 特集 データ主導経済と社会変革. 平成 29 年版 情報通信白書 ICT 白書 2017, 第 1 部, 第 2 章, pp. 79.
- [2] 干川尚人, 下馬場朋禄, 伊藤智義, “地産地消型アーキテクチャによるセンサネットワークデータのプライバシー保護,” 情報処理学会論文誌, vol.59, No.12, pp.2180-2190, Dec, 2018.
- [3] “AWS IoT プラットフォーム” .
<https://aws.amazon.com/jp/s/dm/optimization/sem/iot-signup/> (参照 2018-12-20)
- [4] “ PaaS で 誰 も が デ ー タ 分 析 ” . 日 経 コ ン ピ ュ ー タ .
<http://tech.nikkeibp.co.jp/it/atclact/active/15/010700161/010700003/> (参照 2018-12-20)
- [5] 干川先生の奴？
- [6] 社会基盤情報流通推進協議会, 人流解析チーム, “松江駅構内人流センサデータ”
- [7] “株式会社アルコム RD-CA213 AHD カメラ 220 万画素バレット型赤外線搭載屋外カメラ” <https://www.arucom.ne.jp/store/g/g250-011409/> (参照 2018-12-20)
- [8] “株式会社アルコム 人相確認距離から防犯カメラを選ぶ。”
<https://www.arucom.ne.jp/sample/> (参照 2018-12-20)
- [9] 関東交通広告協議会, “交通広告調査レポート 2009,” 鉄道利用者プロフィール, pp39